DOCUMENT RESUME

ED 064 393 TM 001 641

AUTHOR	Rogers, W. Todd
TITLE	Jackknifing Disattenuated Correlations: National Assessment of Educational Progress.
INSTITUTION	Colorado Univ., Boulder. Lab. of Educational Research.
PUB DATE	[72]
NOTE	52p.
AVAILABLE FROM	Laboratory of Educational Research University of Colorado, Boulder, Colorado
EDRS PRICE	MF-$0.65 HC Not Available from EDRS.
DESCRIPTORS	Computer Science; *Correlation; *Educational Programs; *Hypothesis Testing; *Models; Problem Solving; *Program Evaluation; Sampling; Simulation; Theories; Values
IDENTIFIERS	*Jackknifing Technique

ABSTRACT

The jackknife is a general inferential technique intended to ameliorate the problems associated with inadequate sampling theory. The research reported herein is directed at investigating the utility of the jackknife for establishing confidence intervals on and testing hypotheses about the disattenuated correlation coefficient for small samples. A review of the literature is made. Several computer simulations were performed to investigate the utility of the jackknife. The theory of the jackknife implies that the jackknife statistic is approximately distributed as a Student - t variate with the appropriate degrees of freedom. Results include: (1) The direction of the difference between the theoretical and actual cumulative proportions of jackknife statistics which were at or below the 10 percentile points of comparison varied across the different values of p(TX, TY); (2) The jackknife was sensitive to changes in the values of the reliabilities for each combination of p(TX, TY) and N;. Conclusions include: (1) It is unlikely that a mathematical model can be formulated to describe the sampling distribution; (2) The performance of the jackknife was sensitive to changes in the values of the input a parameters; and (3) The jackknife can be used to set approximate confidence intervals. (CK)

5.10

ED 064393

RESEARCH PAPERS

Laboratory of Educational Research

UNIVERSITY OF COLORADO

TM 001 641

Not to be cited
without permission

Jackknifing Disattenuated Correlations*

W. Todd Rogers

National Assessment of Educational Progress

Despite authoritative support for the use of disattenuated correlations in certain experimental situations, their actual use in these situations has been limited. This may partially be explained by the restricted availability of inferential procedures due to the lack of adequate distribution theory for the disattenuated correlation coefficient. The jackknife (Tukey, 1958) is a general inferential technique intended to ameliorate the problems associated with inadequate sampling theory. The research reported herein is directed at investigating the utility of the jackknife for establishing confidence intervals on and testing hypotheses about the disattenuated correlation coefficient for small samples.

Disattenuated Correlation Coefficients

Classical test theory (Gulliksen, 1950; Lord and Novick, 1968) is based upon the assumption that an observed score for an examinee can be regarded as the sum of two unobservable components: a true T and an error of measurement E. It is assumed that the error score for one measurement is independent of that for another.

---

* This paper is based upon part of the author's doctoral dissertation (Rogers, 1971) completed at the University of Colorado. The author is greatly indebted to Dr. Gene V Glass, who, as thesis advisor, generously contributed his time, guidance and encouragement.

Consequently, if it can be assumed that the relationship between two sets of measurements is linear, the Pearson-product moment correlation between two sets of observed scores will be lower than what it would have been had the measurements been error free. Spearman (1904a) called this lowering of a correlation coefficient due to the fallibility of measurement <u>attenuation</u> and provided the basic procedure for estimating what the value of a correlation coefficient would be if the errors of measurement were eliminated. The basic formula is:

$$\rho(T_X,T_Y) = \frac{\rho(X,Y)}{\{\rho(X,X')\ \rho(Y,Y')\}^{\frac{1}{2}}} \quad , \qquad (1)$$

where $\rho(X,Y)$ is the population correlation between observed scores on test X and test Y, $\rho(X,X')$ and $\rho(Y,Y')$ are the population values of the reliabilities of test X and test Y, and $\rho(T_X,T_Y)$ is the correlation between true scores for test X and test Y. Lord (1957) called the corrected correlation the <u>disattenuated</u> <u>correlation</u>.

Sample estimates, $r(T_X,T_Y)$, of $\rho(T_X,T_Y)$ are obtained by substituting in formula (1) sample estimates for $\rho(X,Y)$, $\rho(X,X')$, and $\rho(Y,Y')$ which incorporate a consistent definition of error, i.e., the reliability estimates treat as error those factors which attenuate the correlation between X and Y. The testing paradigm simulated in the present study involves administering both test X and test Y on one test occasion followed by a second administration after a suitable time has elapsed:

X_1 X_2

Y_1 Y_2

An estimate of $\rho(T_X, T_Y)$ consistent with this design is:

$$r(T_X, T_Y) = \frac{r(X_1, Y_2) + r(X_2, Y_1)}{2\{r(X_1, X_2)\ r(Y_1, Y_2)\}} \quad (2)$$

Several other formulas for obtaining estimates of $\rho(T_X, T_Y)$ consistent with this design have been presented in the literature (Spearman, 1904a, 1907; Yule (see Appendix C in Spearman, 1910); Lord, 1957). Formula (2) differs from Yule's formula in that the arithmetic mean of the correlations between test X and test Y is used instead of their geometric mean. If it can be assumed that the series of measurements have been conducted with equal accuracy, then the differences in value between $r(X_1, Y_2)$ and $r(X_2, Y_1)$ can be attributed to chance and can, therefore, be taken care of by taking an average. In contrast to the formulas of Spearman and Lord, the same administration correlations between test X and test Y ($r(X_1, Y_1)$ and $r(X_2, Y_2)$) are ignored since the factor of contemporaneity, which can enhance these correlations, is not present in the estimates of reliability.

Both the correlations between repeated measures (test-retest reliabilities) and the correlations between distinct measures include as error variance fluctuations from one test occasion to the next. In contrast, internal analyses and parallel forms estimates of reliability generally result in an inconsistent definition of error since specificity, which is treated as error in these reliability estimates, does not lower the correlation between $\underline{X}$ and $\underline{Y}$ (Johnson, 1950).

Despite the availability of sample estimators and authoritative support for the use of disattenuated correlations in certain situations (e.g., Thouless, 1939; Gulliksen, 1950; Block, 1963; Lord, 1957, 1970; Lord and Novick, 1968; Cochran, 1970), their actual use has been limited. This may partially be explained by the restricted availability of inferential procedures due to the lack of an adequate distribution theory for the disattenuated correlation coefficient. In contrast to the uncorrected coefficient, the theoretical distribution of $\underline{r}(\underline{T}_X,\underline{T}_Y)$ has not yet been derived; in view of the complexity of the sample estimator, a theoretical derivation of its sampling distribution probably leads to intractable mathematics, making an exact analytical solution exceedingly difficult, if not impossible. Formulas for obtaining approximate values of the standard error corresponding to the various sample estimators of $\underline{\rho}(\underline{T}_X,\underline{T}_Y)$ have been derived by Kelley (1923, 1947), Shen (1924), and Cureton and Dunlap (1930). The formula for the standard error of $\underline{r}(\underline{T}_X,\underline{T}_Y)$ computed from formula (2) is:

$$\text{S.E.}\{r(T_X,T_Y)\} = \frac{r(T_X,T_Y)}{\{2(N-1)\}^{\frac{1}{2}}}\left[\frac{(1-r^2(X_1,X_2))^2}{2r^2(X_1,X_2)} + \frac{(1-r^2(Y_1,Y_2))^2}{2r^2(Y_1,Y_2)} + \frac{1-r^2(C)+r^2(A)+r(X_1,X_2)r(Y_1,Y_2)}{r^2(C)} - \frac{2r(A)r(C)(r(X_1,X_2)+r(Y_1,Y_2))}{r^2(C)}\right.$$

$$+ \frac{(r^2(A) + r^2(C))(1 + r(X_1,X_2)r(Y_1,Y_2)}{r(X_1,X_2)r(Y_1,Y_2)}$$

$$- \frac{2r(A)r(C)(r(X_1,X_2) + r(Y_1,Y_2))}{r(X_1,X_2)r(Y_1,Y_2)} - \frac{2r(A)}{r(C)} \cdot$$

$$\left[\frac{1 - r^2(X_1,X_2) - r^2(C)}{r(X_1,X_2)} + \frac{1 - r^2(Y_1,Y_2) - r^2(C)}{r(Y_1,Y_2)} \right]$$

$$+ \frac{2r^2(A) - 4r^2(C) - r^2(X_1,X_2) - r^2(Y_1,Y_2)}{2}$$

$$+ 2 - 2r^2(A) \Bigg]^{\frac{1}{2}} , \qquad (3)$$

where

$$r(C) = \frac{r(X_1,Y_2) + r(X_2,Y_1)}{2} ,$$

$$r(A) = \frac{r(X_1,Y_1) + r(X_2,Y_2)}{2} ,$$

and N equals the number of subjects.

Formula (3) illustrates the complexity of the standard error formulas for $r(T_X,T_Y)$. The method of logarithmic differentials used in the derivations of these formulas is a large sample procedure (Kelley, 1947, p. 523). The absence of a theoretical sampling distribution for the disattenuated correlation coefficient and the questionable validity of the standard error formulas for small to moderately sized samples limits the general applicability of these formulas.

Although a derivation of the theoretical sampling distribution of $\underline{r}(\underline{T}_X, \underline{T}_Y)$ has not been obtained, procedures for making tests of hypotheses and constructing confidence intervals in special situations have been developed. Lord (1957) and McNemar (1958) developed two different procedures for testing the hypothesis that $\underline{\rho}(\underline{T}_X, \underline{T}_Y) = 1.0$. Lord's procedure, developed using maximum likelihood procedures, is a large sample test. McNemar's test is based upon analysis of variance. His method, although appropriate for any sample size, assumes that the two tests are equally reliable in the population. Neither of these tests has been generalized to other hypotheses about $\underline{\rho}(\underline{T}_X, \underline{T}_Y)$.

DuBois (1965) suggested that the disattenuated correlation coefficient was equal to the uncorrected correlation between X and Y with the error components partialled out, i.e.,

$$r(T_X, T_Y) = r(X, Y \cdot E_X, E_Y) \ .$$

This method of computing corrected coefficients and describing its distributions in terms of distributions of partial correlations does not seem to be of practical value. In the matrix of correlations DuBois used to determine $\underline{r}(\underline{T}_X, \underline{T}_Y)$, he assumed that $\underline{r}(\underline{E}_X, \underline{E}_Y)$, $\underline{r}(\underline{E}_X, \underline{Y})$, and $\underline{r}(\underline{E}_Y, \underline{X})$ are equal to zero. Although it is assumed in classical test theory that the corresponding population values are zero, there is no reason to expect these sample values to be zero for a particular sample.

The most recent studies of inferential procedures for testing hypotheses about $\underline{\rho}(\underline{T}_X, \underline{T}_Y)$ and establishing confidence limits for $\underline{\rho}(\underline{T}_X, \underline{T}_Y)$ are the Monte Carlo studies performed by Forsyth (1967), and Forsyth and Feldt (1969, 1970). Based on their finding that "for suitably large samples, the sampling distribution of $\underline{r}(\underline{T}_X, \underline{T}_Y)$ is approximately normal in form" (Forsyth and Feldt, 1969, p. 65), Forsyth and Feldt investigated using standard inferential procedures based on normal curve theory. The standard errors were calculated using formulas derived by the method of logarithmic differentials. The data obtained indicated close agreement between the actual proportion of confidence intervals enclosing $\underline{\rho}(\underline{T}_X, \underline{T}_Y)$ and the nominal level for values of α equal to .10 and .05. However, the sizes of the critical regions corresponding to a two-tail hypothesis test were consistently uneven. To overcome the uneven distribution of Type I errors, Forsyth and Feldt used the hypothesized value of $\underline{\rho}(\underline{T}_X, \underline{T}_Y)$ which, in their computer simulation, was actually equal to the parameter value used in generating the sampling distributions. In practice the value of $\underline{\rho}(\underline{T}_X, \underline{T}_Y)$ is not known. To the extent the hypothesized value differs from the actual value of $\underline{\rho}(\underline{T}_X, \underline{T}_Y)$, the Type I errors will be unevenly distributed to an unknown degree. This fact severely restricts the use of normal curve procedures for directional hypothesis testing.

Present inferential procedures are restricted to special situations. The procedures for hypothesis testing developed by Lord and McNemar have not been generalized to other hypotheses about $\underline{\rho}(\underline{T}_X, \underline{T}_Y)$. The normal curve procedure proposed by Forsyth and Forsyth and Feldt appears to be restricted to large samples. Further, the method of logarithmic differentials used to derive the formula for the standard error of $\underline{r}(\underline{T}_X, \underline{T}_Y)$ is a large sample procedure. The jackknife is an inferential procedure intended to obtain approximate confidence intervals simply in problems where standard statistical procedures may not exist or are difficult to apply.

The Jackknife

Statistical inference is the process of generalizing from known characteristics of a sample to the corresponding but unknown characteristics of the population from which the sample was drawn. Given a sample of observations, the inferential procedure involves (a) obtaining a function of these observations which should provide an estimate of the parameter of interest, (b) obtaining a measure of the precision of the estimate, and (c) combining the obtained estimate and the measure of precision with knowledge of the sampling distribution of the estimator to make probabilistic statements about the value of the parameter. There exist situations in which it is possible to determine an estimate of a parameter but yet for which it is difficult, if not impossible, to derive a measure of precision or for which the sampling distribution of the estimator is either not known or else very complex. In these situations one cannot often simply use known standard procedures to obtain an idea of the accuracy of the estimate. The jackknife (Tukey, 1958) is a procedure which may be used in situations of this type to obtain approximate confidence intervals simply in terms of the estimator of the parameter. Tukey adopted the name jackknife since, like a boy scout's jackknife, the procedure is intended to be generally applicable but, like the scout's jackknife, many of its jobs could be better done by a specialized tool, particularly if that tool were available.

The jackknife depends upon dividing a set of data into groups, obtaining estimates from combinations of these groups, and averaging these estimates. Let θ be the unknown parameter, and let $\{X_1, X_2, \ldots, X_N\}$ be a sample of N independent, identically distributed observations with continuous density function F_θ, which depends upon θ. Suppose a method for estimating θ is available. The jackknife requires that the N observations be divided into k ($k \geq 1$) groups of size n ($n \geq 1$) such that $N = nk$, i.e., $(X_1 \ldots, X_n; X_{n+1}, \ldots X_{2n}; \ldots; X_{(k-1)n+1}, \ldots, X_{kn})$. Let $\hat{\theta}$ denote the estimate of θ based on all N observations and let $\hat{\theta}_{-i}$, $i = 1, \cdots, k$, denote the estimate of θ based on the $(N - n)$ observations in the subsample obtained by omitting the ith group.

New estimates of θ, called pseudo-values, are formed by taking a linear combination of $\hat{\theta}$ and the $\hat{\theta}_{-i}$'s:

$$\hat{\theta}_{*i} = k\hat{\theta} - (k - 1)\,\hat{\theta}_{-i}, \quad i = 1, \cdots, k.$$

The jackknife estimate of θ is the mean of the pseudo-values:

$$\hat{\theta}_{*.} = \frac{\sum_{i=1}^{k} \hat{\theta}_{*i}}{k} \quad .$$

The jackknife estimate, $\hat{\theta}_{*.}$, was first introduced by Quenouille (1949, 1956) as a method for reducing bias of the form $\frac{1}{N}$ in parametric estimation. Tukey proposed that in most situations the k

pseudo-values, $\hat{\underline{\theta}}_{*1}, \ldots, \hat{\underline{\theta}}_{*\underline{k}}$, could be treated as $\underline{k}$ approximately independent, identically distributed observations from which an approximate $\underline{t}$ - statistic confidence interval for $\underline{\theta}$ could be constructed. An estimate of the standard error of the jackknife estimate is given by:

$$s_{\hat{\theta}_{*.}} = \left[\frac{\sum_{i=1}^{k} (\hat{\theta}_{*i} - \hat{\theta}_{*.})^2}{k(k-1)}\right]^{\frac{1}{2}}. \qquad (4)$$

Tukey's proposal implies that the quantity

$$\frac{\hat{\theta}_{*.} - \theta}{s_{\hat{\theta}_{*.}}}, \qquad (5)$$

the jackknife statistic (Collins, 1970, p. 53), is approximately distributed as a Student - $\underline{t}$ variate with $\underline{k}$ - 1 degrees of freedom.

> The key idea is that, in a wide variety of problems, the pseudo-values can be used to set approximate confidence limits, using Student's t, as if they were the results of applying some complex calculation to each of $\underline{k}$ independent pieces of data. The words "as if" are vital here; Student's $\underline{t}$ performs well in many circumstances where the θ_{*i} deviate substantially from independence. (Mosteller and Tukey, 1968, p. 135).

A comment is required on the degrees of freedom. The correct number of degrees of freedom for the variance of the jackknife estimate may be less than $\underline{k}$ - 1. For example, jackknifing the median of a sample of size $\underline{N}$ = 2$\underline{m}$, $\underline{m}$ a whole number, by deleting one observation at a time results in only one degree of freedom and not 2$\underline{m}$ - 1. If any one observation is deleted in the upper half of the 2$\underline{m}$ observations ranked according to size, the median of the remaining $2\underline{m} - 1$ observations will be the $\underline{m}$th observation (counting up from the lowest). For any observation deleted from the lower half of the distribution, the median will be the $(\underline{m} + 1)$st value in the ranked distribution of observations. This implies only one degree of freedom for $s_{\hat{\theta}_{*.}}$ (Mosteller and Tukey point out that by jackknifing in groups in which $\underline{n} > 1$, more than two different pseudo-values will occur.) Mosteller and Tukey (1968, p. 136) recommended the following general rules of thumb for determining the proper degrees of freedom for the jackknife estimate:

> c) Count the number of different numbers appearing as pseudo-values, subtract one, and use the results as degrees of freedom.
>
> . . . .
>
> c1) If slight changes in the basic observations--as when values by their nature either 0 or 1 are made -0.001, + .002, 0.997, or 1.004--would make two pseudo-values different, they should not be considered "the same" in applying rule (c). . . .
>
> c2) If carrying more decimals in the computation would have made two pseudo-values different, they should not be considered "the same" in applying rule (c).

Jackknifing Functions of Statistics

Functions of statistics can be divided into two categories: linear and non-linear. For linear functions, the order of operation for obtaining the jackknife estimate does not matter. The jackknife estimate of a linear combination of statistics is equal to the linear combination of the jackknife estimate of each of the statistics that enter into the linear combination, i.e., if

$$\theta' = a_1\hat{\theta}_1 + a_2\hat{\theta}_2 + \ldots + a_m\hat{\theta}_m + b,$$

then $$\hat{\theta}'_{*.} = a_1\hat{\theta}_{1*.} + a_2\hat{\theta}_{2*.} + \ldots + a_m\hat{\theta}_{m*.} + b.$$

For non-linear functions, the order of operations does matter. If $\underline{\hat{\theta}}' = f(\underline{\hat{\theta}}_1, \underline{\hat{\theta}}_2, \ldots, \underline{\hat{\theta}}_m)$ is a non-linear function of the statistics $\underline{\hat{\theta}}_1, \underline{\hat{\theta}}_2, \ldots, \underline{\hat{\theta}}_m$, then

$$\theta'_{*.} \neq f(\hat{\theta}_{1*.}, \hat{\theta}_{2*.}, \ldots, \hat{\theta}_{m*.}).$$

For example, consider the logarithmic function

$$\hat{\theta}' = \log(\hat{\theta}).$$

$$\hat{\theta}'_{*i} = k \log\{\hat{\theta}\} - (k-1)\log\{\hat{\theta}_{-i}\}$$

$$\hat{\theta}'_{*.} = \frac{\sum_{i=1}^{k} \hat{\theta}'_{*i}}{k}$$

$$= \frac{\sum_{i=1}^{k} (k \log\{\hat{\theta}\} - (k-1)\log\{\hat{\theta}_{-i}\})}{k}$$

$$= k \log\{\hat{\theta}\} - (k-1) \frac{\sum_{i=1}^{k} \log\{\theta_{-i}\}}{k}$$

$$= \log\left[\frac{\hat{\theta}^k}{\prod_{i=1}^{k} \hat{\theta}_{-i}^{(k-1)/k}}\right].$$

But

$$\log\{\hat{\theta}_{*.}\} = \log\left[\frac{\sum_{i=1}^{k}\hat{\theta}_{*i}}{k}\right]$$

$$= \log\left[k\hat{\theta} - \frac{(k-1)}{k}\sum_{i=1}^{k}\hat{\theta}_{-i}\right]$$

$$\neq \log\left[\frac{\hat{\theta}^{k}}{\prod_{i=1}^{k}\theta_{-i}^{(k-1)/k}}\right]$$

Therefore,

$$\hat{\theta}'_{*.} \neq \log\{\hat{\theta}_{*}.\}.$$

The above inequality suggests two alternatives for jack-knifing non-linear functions. First, jackknife the entire function:

$$\hat{\theta}_{*.} = \frac{1}{k}\sum_{i=1}^{k}\left[kf(\hat{\theta}_1, \hat{\theta}_2, \ldots\hat{\theta}_m) - (k-1)f(\hat{\theta}_{1_{-i}}, \hat{\theta}_{2_{-i}}, \ldots, \hat{\theta}_{m_{-i}})\right].$$

Second, form the function of the pseudo-values of each term in the combination:

$$\hat{\theta}'_{*i} = f(\hat{\theta}_{1*_i}, \hat{\theta}_{2*_i}, \ldots, \hat{\theta}_{*i}).$$

The mean of these pseudo-values yields the jackknife estimate:

$$\hat{\theta}_{*.} = \frac{\sum_{i=1}^{k} \hat{\theta}_{*i}}{k} .$$

The standard error of the jackknife estimate for both methods is calculated using formula (4).

Miller (1964) gave the conditions under which the jackknife estimate of a function of the sample mean is asymptotically, normally distributed with the correct mean and variance. Arvensen (1969) extended Miller's work to functions of U - statistics and Brillinger proved that the jackknife was asymptotically correct when jackknifing maximum likelihood functions. Examples of the successful application of the jackknife procedure include ratio estimation (Mosteller and Tukey, 1968), variances computed from one dimensional arrays (Miller, 1968) and two dimensional arrays (Collins, 1970) and analysis of sample surveys (Brillinger, 1964; Arvensen, 1969; Frankel, 1971; National Assessment of Educational Progress, 1970). However the jackknife is not foolproof. Difficulties were illustrated when the sampling distribution of the statistic to be jackknifed was asymmetrical or had a straggling tail (Miller,1964). Collins (1970) found that the jackknife performed poorly on the generalizability coefficient (a function of variance components). Mosteller and Tukey (1968) warned of problems when the possible values of the parameter to be estimated are restricted to an interval or half-line. They (Mosteller and Tukey, 1968, pp. 137-138) summarized their discussion on jackknifing statistics as follows:

> There may be some advantage in jackknifing one expression of a given result rather than another (as when we jackknife log y or y^2 instead of y).
> ...
> We know little about which choices of expression tend to polish up the behavior of the jackknife. What evidence we have suggests that:
>
> b1) It is very desirable to avoid situations where the sampling distribution of the quantity jackknife has an abrupt terminus or where the possible values of its estimand are restricted to an interval or half-line. . . .
>
> b2) It is desirable to avoid sampling distributions with one or more straggling tails.
>
> b3) It is probably desirable to avoid markedly unsymmetrical sampling distributions.
>
> In summary, we can use the jackknifing of several numerical results to tell us about any combination of these results. Our conclusions will usually differ somewhat from those reached by jackknifing that combination directly. This offers us choices that sometimes can allow us to improve our conclusions.

The disattenuated correlation coefficient is a non-linear function of the four other correlation coefficients:

$$r(T_X,T_Y) = \frac{r(X_1,Y_2) + r(X_2,Y_1)}{2\{r(X_1,X_2)\,r(Y_1,Y_2)\}^{\frac{1}{2}}}$$

The possible values of $\underline{\rho}(\underline{T_X},\underline{T_Y})$ are restricted to the closed interval $[-1,1]$. Based on the findings of Forsyth and Feldt (1969), the sample disattenuated correlation coefficient for small samples is non-normally distributed. As stated previously, present inferential procedures are limited to special situations.

The successful application of the jackknife in this situation would permit the use of inferential methods where none are presently available.

Method

Several computer simulations were performed to investigate the utility of the jackknife for establishing confidence intervals on and testing hypotheses about the disattenuated correlation coefficient for small samples. Forty-five combinations of $\rho(T_X, T_Y)$ (1.00, .90, .80, .50, and .00), $\rho(X, X')$ and $\rho(Y, Y')$ (.90, .80; .80, .80; .90, .50), and N (15, 30, 60) were included. A weighting of the factors of economy and accuracy resulted in a decision to include 1000 points in each sampling distribution of the jackknife statistic. All runs were made on a Control Data Corporation (CDC) 6400 System provided by the University of Colorado Graduate School Computing Center.

The theory of the jackknife implies that the jackknife statistic is approximately distributed as a Student - t variate with the appropriate degrees of freedom. For most practical problems, the most crucial area of fit is in the tails of the distributions. Therefore, to test the validity of the jackknife, the theoretical and actual cumulative proportions of jackknife statistics which were at or below the .005, .010, .025, .050, .100, .900, .950, .975, .990, .995 percentile points in the Student - t distribution with the k - 1 degrees of freedom were compared. For each run, k = N.

Five initial simulation runs, corresponding to the five values of $\rho(\underline{T}_X,\underline{T}_Y)$, $\rho(\underline{X},\underline{X}') = .90$, $\rho(\underline{Y},\underline{Y}') = .50$, and $\underline{N} = 15$, were performed for $\underline{r}(\underline{T}_X,\underline{T}_Y)$ and for each of six power functions, two logarithmic functions, and one trignometric function (see Table 2, page 22). Forty-five simulations, corresponding to all the combinations of the input parameters, were performed for $\underline{r}(\underline{T}_X,\underline{T}_Y)$ and for the transformation of $\underline{r}(\underline{T}_X,\underline{T}_Y)$ which yielded the best solution for the five initial simulations. The empirical sampling distribution of $\underline{r}(\underline{T}_X,\underline{T}_Y)$ was generated for each of the 45 combinations of input parameters to obtain a description of the statistic jackknifed.

Results

Characteristics of the Sampling Distribution of $\underline{r}(\underline{T}_X,\underline{T}_Y)$

The mean, variance, skewness, and kurtosis for each of the 45 sampling distributions of $\underline{r}(\underline{T}_X,\underline{T}_Y)$ are presented in Table 1. The skewness measure equalled the third moment about the mean divided by the cube of the standard deviation of the sampling distribution of $\underline{r}(\underline{T}_X,\underline{T}_Y)$. The kurtosis equalled the fourth moment about the mean divided by the fourth power of the standard deviation of the sampling distribution of $\underline{r}(\underline{T}_X,\underline{T}_Y)$.

The means of the sampling distributions of $\underline{r}(\underline{T}_X,\underline{T}_Y)$ tended to be lower than the parameter values except for the case in which the reliabilities were most different (viz., $\rho(\underline{X},\underline{X}') =$

TABLE 1

SUMMARY STATISTICS FOR THE SAMPLING DISTRIBUTION OF $r(T_X, T_Y)$

$\rho(T_X,T_Y)$	$\rho(X,X')$	$\rho(Y,Y')$	$\bar{r}(T_X,T_Y)$	$s^2\{r(T_X,T_Y)\}$	Skewness	Kurtosis
N = 15						
1.00	0.90	0.80	1.0041	0.0025	3.8816	48.7556
1.00	.80	.80	1.0111	.0142	10.0077	149.0011
1.00	.90	.50	1.0578	.0918	4.9135	39.9687
0.90	.90	.80	0.8984	.0082	- 1.2706	7.5522
.90	.80	.80	.8936	.0109	- 1.5145	8.8308
.90	.90	.50	.9409	.0836	4.8727	48.2410
.80	.90	.80	.7845	.0209	- 1.7719	8.4960
.80	.80	.80	.7877	.0268	- 1.4811	6.8610
.80	.90	.50	.8126	.0937	3.1731	30.0633
.50	.90	.80	.4923	.0589	- 1.2523	8.2990
.50	.80	.80	.4984	.0624	- 0.5768	4.7328
.50	.90	.50	.4963	.1092	.4567	6.8760
.00	.90	.80	- .0158	.0842	- .1129	2.7533
.00	.80	.80	.0004	.0951	- .2595	3.0299
0.00	0.90	0.50	0.0018	0.1402	- 0.0399	3.4203
N = 30						
1.00	0.90	0.80	1.0001	0.0006	0.4247	4.6997
1.00	.80	.80	1.0040	.0016	.2609	4.6064
1.00	.90	.50	1.0252	.0198	2.5414	18.5875
.90	.90	.80	0.8991	.0031	- 0.9654	4.9369
0.90	0.80	0.80	0.8975	0.0045	- 0.7042	3.9543

TABLE 1 (CONTINUED)

$\rho(T_X,T_Y)$	$\rho(X,X')$	$\rho(Y,Y')$	$\bar{r}(T_X,T_Y)$	$s^2\{r(T_X,T_Y)\}$	Skewness	Kurtosis
0.80	0.80	0.50	0.9162	0.0195	1.7623	13.3890
.80	.90	.80	.7932	.0075	-1.0413	5.1261
.80	.80	.80	.7947	.0095	-1.0196	5.6234
.80	.90	.50	.8161	.0342	4.7485	60.9789
.50	.90	.80	.4979	.0269	-0.6027	3.5621
.50	.80	.80	.4961	.0270	- .5629	3.3547
.50	.90	.50	.5000	.0445	- .0079	4.6842
.00	.90	.80	- .0001	.0424	- .0715	2.7976
.00	.80	.80	- .0027	.0450	- .0668	2.8680
0.00	0.90	0.50	-0.0057	0.0623	-0.1706	3.0380
N = 60						
1.00	0.90	0.80	1.0009	0.0003	0.1730	3.6678
1.00	.80	.80	1.0023	.0007	.0890	3.4029
1.00	.90	.50	1.0088	.0055	.8912	6.5697
0.90	.90	.80	0.8973	.0014	- .6306	3.7275
.90	.80	.80	.8994	.0020	- .5250	3.9822
.90	.90	.50	.9076	.0073	.6473	7.3864
.80	.90	.80	.7978	.0033	- .7076	3.8892
.80	.80	.80	.7973	.0044	- .5430	3.3916
.80	.90	.50	.8034	.0089	- .0852	3.3507
.50	.90	.80	.4933	.0121	- .2998	3.0643
.50	.80	.80	.4947	.0133	- .3506	3.2702
.50	.90	.50	.5068	.0195	- .3660	3.6712
.00	.90	.80	- .0018	.0202	- .0051	2.9021
.00	.80	.80	- .0023	.0213	- .0135	2.7509
0.00	0.90	0.50	-0.0042	0.0282	-0.1483	3.3828

.90, $\rho(Y,Y')$ = .50). In this latter case, there was a general tendency for the mean of the sampling distribution to overestimate $\rho(T_X,T_Y)$. The bias in $r(T_X,T_Y)$ decreased as the sample size increased. For example, with N = 15, the bias was approximately .058, while for N = 60, the bias was less than .009.

The variance of the sampling distributions of $r(T_X,T_Y)$ decreased as the value of $\rho(T_X,T_Y)$ increased for fixed values of the reliabilities and as the value of N increased. The variances ranged between .0025 and .1402 for N = 15, between .0006 and .0623 for N = 30, and between .0003 and .0282 for N = 60. Within each combination of $\rho(T_X,T_Y)$ and N, the largest variance was obtained with $\rho(X,X')$ = .90 and $\rho(Y,Y')$ = .50.

The skewness and kurtosis of the sampling distribution of $r(T_X,T_Y)$ decreased as the value of N increased. For each combination of N and pair or reliabilities, the $r(T_X,T_Y)$ distributions were positively skewed with $\rho(T_X,T_Y)$ = 1.00, but with $\rho(T_X,T_Y)$ = .00, the distributions were negatively skewed. With $\rho(T_X,T_Y)$ between zero and one, the distributions were negatively skewed for $\rho(X,X')$ = .90, $\rho(Y,Y')$ = .80 and for $\rho(X,X')$ = .80, $\rho(Y,Y')$ = .80. For $\rho(X,X')$ = .90, $\rho(Y,Y')$ = .50, the distributions tended to be positively skewed as the value of $\rho(T_X,T_Y)$ increased and as the value of N decreased.

With $\rho(T_X,T_Y)$ greater than zero, the sampling distributions were leptokurtic for each combination of N and pair of reliabilities. In general the kurtosis increased as the value of $\rho(T_X,T_Y)$ increased. Within each combination of N and $\rho(T_X,T_Y)$, the kurotsis was, in general, greatest with $\underline{\rho}(\underline{X},\underline{X}') = .90$, $\underline{\rho}(\underline{Y},\underline{Y}') = .50$.

<u>Findings Related to the Jackknife</u>.

The five initial simulations performed for each of the statistics jackknifed revealed that the performance of the jackknife was not consistent across the five values of $\underline{\rho}(\underline{T}_X,\underline{T}_Y)$. For example, there was closer agreement between the theoretical and actual cumulative proportions of jackknife statistics produced by the jackknife on $\underline{r}(\underline{T}_X,\underline{T}_Y)$ for $\underline{\rho}(\underline{T}_X,\underline{T}_Y) = .00$ (mean absolute difference equalled .0147) than for $\underline{\rho}(\underline{T}_X,\underline{T}_Y) = 1.00$ (mean absolute difference equalled .0249). Furthermore, the performance of the jackknife on the transformations investigated was not consistent. With $\underline{\rho}(\underline{T}_X,\underline{T}_Y) = .90$ the jackknife on $\underline{r}^{9/5}(\underline{T}_X,\underline{T}_Y)$ yielded superior results to the jackknife on $\underline{r}^{7/5}(\underline{T}_X,\underline{T}_Y)$, but with $\underline{\rho}(\underline{T}_X,\underline{T}_Y) = .50$, the superiority was reversed.

The overall mean absolute difference across all five simulations performed for each of the 11 statistics jackknifed is reported in Table 2. Jackknifing $\underline{r}(\underline{T}_X,\underline{T}_Y)$ by parts and jackknifing power functions of $\underline{r}(\underline{T}_X,\underline{T}_Y)$ in which the exponents were less than one or greater than two resulted in noticeably inferior solutions in comparison to the solution on $\underline{r}(\underline{T}_X,\underline{T}_Y)$. The jackknife results obtained for the remaining transformations were reasonably comparable. The mean of the absolute differences between the theoretical and actual cumulative proportions of jackknife statistics at the 10 points of comparison across all five simulations ranged between .0160 and .0219, with the best solution obtained from the jackknife on $\underline{r}^{7/5}(\underline{T}_X,\underline{T}_Y)$.

Simulations for each of the 45 combinations of $\underline{\rho}(\underline{T}_X,\underline{T}_Y)$, $\underline{\rho}(X,X')$, $\underline{\rho}(Y,Y')$, and N were performed for the jackknife on

TABLE 2

OVERALL MEAN ABSOLUTE DIFFERENCES OBTAINED FROM FIVE INITIAL SIMULATIONS

Statistic	Overall Mean Absolute Difference[a]
$r(T_X,T_Y)$	0.0188
$r^{[illegible]/5}(T_X,T_Y)$	.0160
$r^{9/5}(T_X,T_Y)$	.0207
$r^{7/3}(T_X,T_Y)$	.0295
$r^{3}(T_X,T_Y)$	.0380
$r^{3/8}(T_X,T_Y)$	.0620
$r^{1/5}(T_X,T_Y)$	.0826
$\log_{10}\{r(T_X,T_Y) + c_1\}$[b]	.0219
$\frac{1}{2}\ln\left[\frac{c_1 + r(T_X,T_Y)}{c_2 - r(T_X,T_Y)}\right]$[b]	.0196
$2 \sin \frac{r(T_X,T_Y)}{RMAX}$[b]	.0179
$r(T_X,T_Y)$ by parts	0.2364

[a]The overall mean absolute difference equals the mean of the absolute differences at the .005, .010, .025, .050, .100, .900, .950, .975, .990, and .995 percentile points of the _t_-distribution with 14 degrees of freedom summed across the five initial simulations.

[b]c_1 = .00001 plus the absolute value of the minimum of $r(T_X,T_Y)$ and the jackknife estimate of $\rho(T_X,T_Y)$ for each simulation performed.
c_2 = .00001 + RMAX.
RMAX = maximum of $r(T_X,T_Y)$ and the jackknife estimate of $\rho(T_X,T_Y)$ for each simulation performed.

$\underline{r}^{7/5}(\underline{T}_X,\underline{T}_Y)$ and $\underline{r}(\underline{T}_X,\underline{T}_Y)$. The jackknife on $\underline{r}(\underline{T}_X,\underline{T}_Y)$ was investigated further to see whether for other values of the input parameters the jackknife results would converge to the results obtained from jackknifing $\underline{r}^{7/5}(\underline{T}_X,\underline{T}_Y)$. The results of these simulations are reported in Tables 3-17. Each table consists of a comparison between the theoretical and empirical cumulative proportions of jackknife statistics at 10 points of the $\underline{t}$-distribution with the correct degrees of freedom. To facilitate the comparison between the two solutions, the corresponding results for $\underline{r}(\underline{T}_X,\underline{T}_Y)$ and $\underline{r}^{7/5}(\underline{T}_X,\underline{T}_Y)$ are presented in each table. In each table, column one contains the theoretical cumulative proportion $\underline{p}$ at each of the 10 points. Column two indicates the proportion of jackknife statistics $\underline{p}$ which were at or below each percentile point for $\underline{r}(\underline{T}_X,\underline{T}_Y)$. Column three contains the difference in cumulative proportion (theoretical minus actual) at each of the 10 points. Columns four and five contain the corresponding information as in columns two and three for $\underline{r}^{7/5}(\underline{T}_X,\underline{T}_Y)$. For example, from Table 17, the actual proportion of jackknife statistics at or below the 2.5th percentile point with $\underline{N}$ = 60 is .033 for $\underline{r}(\underline{T}_X,\underline{T}_Y)$. In this case, the actual proportion exceeds the theoretical proportion by .008. For $\underline{r}^{7/5}(\underline{T}_X,\underline{T}_Y)$, the actual proportion of jackknife statistics at or below the 2.5th percentile point with $\underline{N}$ = 60 is .007. For this case, the theoretical proportion exceeds the actual proportion by .018. The means of the absolute differences between the theoretical and actual cumulative proportions at the .005, .010, .025, .050, .100, .900,

TABLE 3

COMPARISON OF ACTUAL ($\hat{p}$) AND THEORETICAL (p) CUMULATIVE PROPORTIONS AT 10 POINTS OF THE t-DISTRIBUTION

INPUT PARAMETERS: $\rho(T_X . T_Y) = 1.00$

$\rho(X,X') = .9$[illegible]

$\rho(Y,Y') = .80$

p	$\hat{p}$ $r(T_X,T_Y)$	$p - \hat{p}$ $r(T_X,T_Y)$	$\hat{p}$ $r^{7/5}(T_X,T_Y)$	$p - \hat{p}$ $r^{7/5}(T_X,T_Y)$
N = 15, d.f. = 14				
0.005	0.003	0.002	0.003	0.002
.010	.004	.006	.004	.006
.025	.010	.015	.011	.014
.050	.021	.029	.022	.028
.100	.056	.044	.059	.041
.900	.968	- .068	.968	- .06[illegible]
.950	.990	- .040	.991	- .041
.975	.994	- .019	.995	.020
.990	.998	- .008	.998	- .008
0.995	0.999	-0.004	0.999	-0.004
N = 30, d.f. = 29				
0.005	0.000	0.008	0.000	0.005
.010	.001	.009	.001	.009
.025	.016	.009	.016	.009
.050	.030	.020	.032	.018
.100	.086	.014	.087	.013
.900	.951	- .051	.953	- .053
.950	.985	- .035	.987	- .024
.975	.999	- .024	.999	- .024
990	1.000	- .010	1.000	- .010
0.995	1.000	-0.005	1.000	-0.005
N = 60, d.f. = 59				
0.005	0.000	0.005	0.000	0.005
.010	.001	.009	.002	.008
.025	.014	.011	.014	.011
.050	.035	.015	.036	.014
.100	.093	.007	.095	.005
.900	.921	- .021	.923	- .023
.950	.970	- .020	.971	- .021
.975	.991	- .016	.991	- .016
.990	.997	- .007	.997	- .007
0.995	0.998	-0.003	0.998	-0.003

TABLE 4

COMPARISON OF ACTUAL ($\hat{p}$) AND THEORETICAL (p) CUMULATIVE PROPORTIONS AT 10 POINTS OF THE t-DISTRIBUTION

INPUT PARAMETERS: $\rho(T_X, T_Y) = 1.00$

$\rho(X, X') = .80$

$\rho(Y, Y') = .80$

p	$\hat{p}$ $r(T_X, T_Y)$	$p - \hat{p}$ $r(T_X, T_Y)$	$\hat{p}$ $r^{7/5}(T_X, T_Y)$	$p - \hat{p}$ $r^{7/5}(T_X, T_Y)$
N = 15, d.f. = 14				
0.005	0.002	0.003	0.002	0.003
.010	.002	.008	.002	.008
.025	.007	.018	.009	.016
.050	.018	.032	.020	.030
.100	.050	.050	.051	.049
.900	.968	- .068	.970	- .070
.950	.993	- .043	.993	- .043
.975	.997	- .022	.497	- .022
.990	.999	- .009	.999	- .009
0.995	0.999	-0.004	0.999	-0.004
N = 30, d.f. = 29				
0.005	0.001	0.004	0.001	0.00[illegible]
.010	.002	.008	.002	.008
.025	.008	.017	.010	.015
.050	.022	.028	.022	.028
.100	.074	.026	.079	.021
.900	.946	- .046	.949	- .049
.950	.988	- .038	.988	- .038
.975	.999	- .024	1.000	- .025
.990	1.000	- .010	1.000	- .010
0.995	1.000	-0.005	1.000	-0.005
N = 60, d.f. = 59				
0.005	0.000	0.005	0.000	0.005
.010	.002	.008	.002	.008
.025	.017	.008	.018	.007
.050	.039	.011	.039	.011
.100	.089	.011	.092	.008
.900	.923	- .023	.927	- .027
.950	.977	- .027	.978	- .028
.975	.993	- .018	.993	- .018
.990	.999	- .009	.999	- .009
0.995	1.000	-0.005	1.000	-0.005

TABLE 5

COMPARISON OF ACTUAL ($\hat{p}$) AND THEORETICAL (p) CUMULATIVE PROPORTIONS AT 10 POINTS OF THE t-DISTRIBUTION

INPUT PARAMETERS: $\rho(T_X,T_Y) = 1.00$

$\rho(X,X') = .90$

$\rho(Y,Y') = .50$

p	$\hat{p}$ $r(T_X,T_Y)$	$p - \hat{p}$ $r(T_X,T_Y)$	$\hat{p}$ $r^{7/5}(T_X,T_Y)$	$p - \hat{p}$ $r^{7/5}(T_X,T_Y)$
N = 15, d.f. = 14				
0.005	0.000	0.005	0.001	0.004
.010	.001	.004	.002	.008
.025	.013	.012	.020	.005
.050	.035	.015	.045	.005
.100	.078	.022	.092	.008
.900	.996	- .096	.996	- .096
.950	1.000	- .050	1.000	- .050
.975	1.000	- .025	1.000	- .025
.990	1.000	- .010	1.000	- .010
0.995	1.000	-0.005	1.000	-0.005
N = 30, d.f. = 29				
0.005	0.000	0.005	0.000	0.005
.010	.004	.006	.004	.006
.025	.011	.014	.017	.008
.050	.040	.010	.049	.001
.100	.094	.006	.106	- .006
.900	.994	- .094	.995	- .095
.950	.999	- .049	1.000	- .050
.975	1.000	- .025	1.000	- .025
.990	1.000	- .010	1.000	- .010
0.995	1.000	-0.005	1.000	-0.005
N = 60, d.f. = 59				
0.005	0.003	0.002	0.004	0.001
.010	.006	.004	.013	- .003
.025	.028	- .003	.031	- .006
.050	.060	- .010	.062	- .012
.100	.117	- .017	.122	- .022
.900	.982	- .082	.984	- .084
.950	.999	- .049	.999	- .049
.975	1.000	- .025	1.000	- .025
.990	1.000	- .010	1.000	- .010
0.995	1.000	-0.005	1.000	-0.005

TABLE 6

COMPARISON OF ACTUAL ($\hat{p}$) AND THEORETICAL (p) CUMULATIVE PROPORTIONS AT 10 POINTS OF THE t-DISTRIBUTION

INPUT PARAMETERS: $\rho(T_X, T_Y) = .90$

$\rho(X, X') = .90$

$\rho(Y, Y') = .80$

p	$\hat{p}$ $r(T_X, T_Y)$	$p - \hat{p}$ $r(T_X, T_Y)$	$\hat{p}$ $r^{7/5}(T_X, T_Y)$	$p - \hat{p}$ $r^{7/5}(T_X, T_Y)$
N = 15, d.f. = 14				
0.005	0.000	0.005	0.001	0.004
.010	.001	.009	.001	.009
.025	.005	.020	.006	.019
.050	.011	.039	.017	.033
.100	.030	.070	.036	.064
.900	.807	.093	.813	.087
.950	.877	.073	.880	.070
.975	.913	.062	.916	.059
.990	.943	.047	.947	.043
0.995	0.964	0.031	0.966	0.029
N = 30, d.f. = 29				
0.005	0.001	0.004	0.001	0.004
.010	.001	.009	.002	.008
.025	.003	.022	.006	.019
.050	.009	.041	.011	.039
.100	.035	.065	.038	.062
.900	.815	.085	.816	.084
.950	.874	.076	.880	.070
.975	.917	.058	.920	.055
.990	.946	.044	.947	.043
0.995	0.959	0.036	0.961	0.034
N = 60, d.f. = 59				
0.005	0.000	0.005	0.000	0.005
.010	.000	.010	.002	.008
.025	.011	.014	.013	.012
.050	.022	.028	.025	.025
.100	.062	.038	.066	.034
.900	.873	.027	.874	.026
.950	.919	.031	.922	.028
.975	.943	.032	.945	.030
.990	.968	.022	.969	.021
0.995	0.975	0.020	0.976	0.019

TABLE 7

COMPARISON OF ACTUAL ($\hat{p}$) AND THEORETICAL (p) CUMULATIVE PROPORTIONS AT 10 POINTS OF THE t-DISTRIBUTION

INPUT PARAMETERS: $\rho(T_X,T_Y) = .90$

$\rho(X,X') = .80$

$\rho(Y,Y') = .80$

p	$\hat{p}$ $r(T_X,T_Y)$	$p - \hat{p}$ $r(T_X,T_Y)$	$\hat{p}$ $r^{7/5}(T_X,T_Y)$	$p - \hat{p}$ $r^{7/5}(T_X,T_Y)$
N = 15, d.f. = 14				
0.005	0.002	0.003	0.003	0.002
.010	.002	.008	.003	.007
.025	.004	.021	.007	.018
.050	.010	.040	.016	.034
.100	.035	.065	.048	.052
.900	.860	.040	.865	.035
.950	.912	.038	.918	.032
.975	.943	.032	.947	.028
.990	.964	.026	.967	.023
0.995	0.972	0.023	0.973	0.022
N = 30, d.f. = 29				
0.005	0.000	0.005	0.000	0.005
.010	.001	.009	.001	.009
.025	.002	.023	.003	.022
.050	.006	.044	.011	.039
.100	.039	.061	.042	.058
.900	.839	.061	.846	.054
.950	.902	.048	.902	.048
.975	.940	.035	.943	.032
.990	.964	.026	.967	.023
0.995	0.976	0.019	0.979	0.016
N = 60, d.f. = 59				
0.005	0.001	0.004	0.001	0.004
.010	.001	.009	.002	.008
.025	.008	.017	.008	.017
.050	.020	.030	.020	.030
.100	.062	.038	.066	.034
.900	.867	.033	.868	.032
.950	.919	.031	.923	.027
.975	.949	.026	.949	.026
.990	.967	.023	.069	.023
0.995	0.976	0.019	0.978	0.017

TABLE 8

COMPARISON OF ACTUAL ($\hat{p}$) AND THEORETICAL (p) CUMULATIVE PROPORTIONS AT 10 POINTS OF THE t-DISTRIBUTION

INPUT PARAMETERS: $\rho(T_X,T_Y) = .90$

$\rho(X,X') = .90$

$\rho(Y,Y') = .50$

p	$\hat{p}$ $r(T_X,T_Y)$	$p - \hat{p}$ $r(T_X,T_Y)$	$\hat{p}$ $r^{7/5}(T_X,T_Y)$	$p - \hat{p}$ $r^{7/5}(T_X,T_Y)$
N = 15, d.f. = 14				
0.005	0.001	0.004	0.003	0.002
.010	.002	.007	.007	.003
.025	.009	.016	.014	.011
.050	.016	.034	.023	.027
.100	.047	.053	.067	.033
.900	.960	- .060	.961	- .061
.950	.979	- .029	.980	- .030
.975	.989	- .014	.991	- .016
.990	.993	- .003	.993	- .003
0.995	0.996	-0.001	0.997	-0.002
N = 30, d.f. = 29				
0.005	0.000	0.005	0.000	0.005
.010	.000	.010	.003	.007
.025	.009	.016	.012	.013
.050	.019	.031	.025	.025
.100	.057	.043	.077	.023
.900	.946	- .046	.954	- .054
.950	.976	- .026	.983	- .033
.975	.992	- .017	.992	- .017
.990	.998	- .008	.998	- .008
0.995	0.998	-0.003	0.999	-0.004
N = 60, d.f. = 59				
0.005	0.001	0.004	0.002	0.003
.010	.002	.008	.006	.004
.025	.008	.017	.010	.015
.050	.028	.022	.036	.014
.100	.068	.032	.076	.024
.900	.917	- .017	.923	- .023
.950	.969	- .019	.971	- .021
.975	.983	- .008	.986	- [illegible]
.990	.995	- .005	.997	- .007
0.995	0.998	-0.003	0.999	-0.004

TABLE 9

COMPARISON OF ACTUAL ($\hat{p}$) AND THEORETICAL (p) CUMULATIVE PROPORTIONS AT 10 POINTS OF THE *t*-DISTRIBUTION

INPUT PARAMETERS: $\rho(T_X,T_Y) = .80$

$\rho(X,X') = .90$

$\rho(Y,Y') = .80$

p	$\hat{p}$ $r(T_X,T_Y)$	$p - \hat{p}$ $r(T_X,T_Y)$	$\hat{p}$ $r^{7/5}(T_X,T_Y)$	$p - \hat{p}$ $r^{7/5}(T_X,T_Y)$
N = 15, d.f. = 14				
0.005	0.000	0.005	0.005	0.000
.010	.002	.008	.008	.002
.025	.008	.017	.018	.007
.050	.017	.033	.031	.019
.100	.048	.052	.067	.033
.900	.809	.091	.816	.084
.950	.880	.070	.883	.067
.975	.915	.060	.919	.056
.990	.939	.051	.945	.045
0.995	0.951	0.044	0.952	0.043
N = 30, d.f. = 29				
0.005	0.001	0.004	0.001	0.004
.010	.001	.009	.003	.007
.025	.005	.020	.012	.013
.050	.015	.035	.020	.030
.100	.039	.061	.049	.051
.900	.852	.048	.858	.042
.950	.900	.050	.901	.049
.975	.924	.051	.929	.046
.990	.948	.042	.952	.038
0.995	0.964	0.031	0.969	0.026
N = 60, d.f. = 59				
0.005	0.000	0.005	0.002	0.003
.010	.003	.007	.004	.006
.025	.007	.018	.010	.015
.050	.019	.031	.023	.027
.100	.053	.047	.059	.041
.900	.855	.045	.858	.042
.950	.914	.036	.922	.028
.975	.951	.024	.952	.023
.990	.990	.017	.975	.015
0.995	0.995	0.017	0.979	0.016

TABLE 10

COMPARISON OF ACTUAL ($\hat{p}$) AND THEORETICAL (p) CUMULATIVE PROPORTIONS AT 10 POINTS OF THE *t*-DISTRIBUTION

INPUT PARAMETERS: $\rho(T_X,T_Y) = .80$

$\rho(X,X') = .80$

$\rho(Y,Y') = .80$

p	$\hat{p}$ $r(T_X,T_Y)$	$p - \hat{p}$ $r(T_X,T_Y)$	$\hat{p}$ $r^{7/5}(T_X,T_Y)$	$p - \hat{p}$ $r^{7/5}(T_X,T_Y)$
N = 15, d.f. = 14				
0.005	0.003	0.002	0.001	0.004
.010	.004	.006	.007	.003
.025	.008	.017	.017	.008
.050	.019	.031	.038	.012
.100	.046	.054	.073	.027
.900	.822	.078	.830	.007
.950	.880	.070	.885	.065
.975	.911	.064	.916	.059
.990	.941	.049	.944	.046
0.995	0.953	0.042	0.956	0.039
N = 30, d.f. = 29				
0.005	0.002	0.003	0.004	0.001
.010	.003	.007	.007	.003
.025	.007	.018	.009	.016
.050	.015	.035	.020	.030
.100	.040	.060	.053	.047
.900	.833	.067	.839	.061
.950	.882	.068	.892	.058
.975	.926	.049	.933	.042
.990	.947	.043	.954	.036
0.995	0.965	0.030	0.970	0.025
N = 60, d.f. = 59				
0.005	0.000	0.005	0.000	0.005
.010	.000	.010	.000	.010
.025	.005	.020	.012	.013
.050	.025	.025	.033	.017
.100	.060	.040	.068	.032
.900	.851	.049	.859	.041
.950	.915	.035	.920	.030
.975	.944	.031	.948	.027
.990	.962	.028	.965	.028
0.995	0.974	0.021	0.976	0.019

TABLE 11

COMPARISON OF ACTUAL ($\hat{p}$) AND THEORETICAL (p) CUMULATIVE PROPORTIONS AT 10 POINTS OF THE t-DISTRIBUTION

INPUT PARAMETERS: $\rho(T_X,T_Y) = .80$

$\rho(X,X') = .90$

$\rho(Y,Y') = .50$

p	$\hat{p}$ $r(T_X,T_Y)$	$p - \hat{p}$ $r(T_X,T_Y)$	$\hat{p}$ $r^{7/5}(T_X,T_Y)$	$p - \hat{p}$ $r^{7/5}(T_X,T_Y)$
N = 15, d.f. = 14				
0.005	0.002	0.003	0.008	-0.003
.010	.003	.007	.012	- .002
.025	.009	.016	.019	.006
.050	.014	.036	.040	.010
.100	.050	.050	.091	.009
.900	.925	- .025	.931	- .031
.950	.957	- .007	.961	- .011
.975	.979	- .004	.982	- .007
.990	.989	.001	.989	.001
0.995	0.989	0.006	0.990	0.005
N = 30, d.f. = 29				
0.005	0.000	0.005	0.004	0.001
.010	.003	.007	.006	.004
.025	.007	.018	.012	.013
.050	.013	.037	.027	.023
.100	.052	.048	.066	.034
.900	.899	.001	.910	- .010
.950	.947	.003	.955	- .005
.975	.971	.004	.977	- .002
.990	.984	.006	.986	.004
0.995	0.990	0.005	0.992	0.003
N = 60, d.f. = 59				
0.005	0.000	0.005	0.001	0.004
.010	.001	.009	.003	.007
.025	.011	.014	.012	.013
.050	.024	.026	.032	.018
.100	.066	.034	.083	.017
.900	.886	.014	.892	.008
.950	.932	.018	.943	.007
.975	.962	.013	.967	.008
.990	.982	.008	.984	.006
0.995	0.986	0.009	0.989	0.006

TABLE 12

COMPARISON OF ACTUAL ($\hat{p}$) AND THEORETICAL (p) CUMULATIVE PROPORTIONS AT 10 POINTS OF THE t-DISTRIBUTION

INPUT PARAMETERS: $\rho(T_X,T_Y) = .50$

$\rho(X,X') = .90$

$\rho(Y,Y') = .80$

p	$\frac{\hat{p}}{r(T_X,T_Y)}$	$\frac{p - \hat{p}}{r(T_X,T_Y)}$	$\frac{\hat{p}}{r^{7/5}(T_X,T_Y)}$	$\frac{p - \hat{p}}{r^{7/5}(T_X,T_Y)}$
N = 15, d.f. = 14				
0.005	0.001	0.004	0.009	-0.004
.010	.004	.006	.017	- .007
.025	.011	.014	.045	- .020
.050	.023	.027	.073	- .023
.100	.058	.042	.123	- .023
.900	.818	.082	.838	.062
.950	.874	.076	.888	.062
.975	.908	.067	.923	.052
.990	.939	.051	.957	.033
0.995	0.961	0.034	0.973	0.022
N = 30, d.f. = 29				
0.005	0.000	0.005	0.023	-0.018
.010	.003	.007	.032	- .022
.025	.011	.014	.051	- .026
.050	.030	.020	.069	- .019
.100	.069	.031	.108	- .008
.900	.831	.069	.844	.056
.950	.884	.066	.902	.048
.975	.923	.052	.937	.038
.990	.952	.038	.964	.026
0.995	0.965	0.030	0.973	0.022
N = 60, d.f. = 59				
0.005	0.004	0.001	0.012	-0.007
.010	.006	.004	.017	- .007
.025	.012	.013	.029	- .004
.050	.027	.023	.051	- .001
.100	.076	.024	.110	- .010
.900	.873	.027	.883	.017
.950	.920	.030	.932	.018
.975	.958	.017	.964	.011
.900	.971	.019	.980	.010
0.995	0.982	0.013	0.985	0.010

TABLE 13

COMPARISON OF ACTUAL ($\hat{p}$) AND THEORETICAL (p) CUMULATIVE PROPORTIONS AT 10 POINTS OF THE t-DISTRIBUTION

INPUT PARAMETERS: $\rho(T_X,T_Y) = .50$

$\rho(X,X') = .50$

$\rho(Y,Y') = .80$

p	$\hat{p}$ $r(T_X,T_Y)$	$p - \hat{p}$ $r(T_X,T_Y)$	$\hat{p}$ $r^{7/5}(T_X,T_Y)$	$p - \hat{p}$ $r^{7/5}(T_X,T_Y)$
N = 15, d.f. = 14				
0.005	0.006	0.001	0.012	-0.007
.010	.009	.001	.025	- .015
.025	.021	.004	.047	- .022
.050	.033	.017	.068	- .018
.100	.054	.046	.109	- .009
.900	.835	.065	.858	.042
.950	.887	.063	.896	.054
.975	.915	.060	.928	.047
.990	.934	.056	.945	.045
0.995	0.949	0.046	0.964	0.031
N = 30, d.f. = 29				
0.005	0.002	0.003	0.031	-0.026
.010	.006	.004	.035	- .028
.025	.017	.008	.047	- .022
.050	.029	.021	.066	- .016
.100	.069	.031	.105	- .005
.900	.851	.049	.871	.029
.950	.908	.042	.920	.030
.975	.933	.042	.953	.022
.990	.964	.026	.972	.018
0.995	0.973	0.022	0.979	0.016
N = 60, d.f. = 59				
0.005	0.004	0.001	0.012	-0.007
.010	.005	.005	.017	- .007
.025	.011	.014	.028	- .003
.050	.030	.020	.064	- .014
.100	.070	.030	.096	- .004
.900	.867	.033	.880	.020
.950	.921	.029	.928	.022
.975	.947	.028	.961	.014
.990	.969	.021	.974	.016
0.995	0.976	0.019	0.982	0.013

TABLE 14

COMPARISON OF ACTUAL ($\hat{p}$) AND THEORETICAL (p) CUMULATIVE PROPORTIONS AT 10 POINTS OF THE *t*-DISTRIBUTION

INPUT PARAMETERS: $\rho(T_X,T_Y) = .50$

$\rho(X,X') = .90$

$\rho(Y,Y') = .50$

p	$\hat{p}$ $r(T_X,T_Y)$	$p - \hat{p}$ $r(T_X,T_Y)$	$\hat{p}$ $r^{7/5}(T_X,T_Y)$	$p - \hat{p}$ $r^{7/5}(T_X,T_Y)$
N = 15, d.f. = 14				
0.005	0.007	-0.002	0.014	-0.009
.010	.010	.000	.019	- .009
.025	.016	.009	.043	- .018
.050	.025	.025	.066	- .016
.100	.052	.048	.117	- .017
.900	.895	.005	.915	- .015
.950	.934	.016	.942	.008
.975	.953	.022	.960	.015
.990	.971	.019	.980	.010
0.995	0.980	0.015	0.981	0.014
N = 30, d.f. = 29				
0.005	0.003	0.002	0.023	-0.018
.010	.003	.007	.035	- .025
.025	.012	.013	.056	- .031
.050	.025	.025	.078	- .028
.100	.057	.043	.127	- .027
.900	.891	.009	.908	- .008
.950	.930	.020	.943	.007
.975	.956	.019	.967	.008
.990	.974	.016	.977	.013
0.995	0.978	0.017	0.987	0.008
N = 60, d.f. = 59				
0.005	0.003	0.002	0.020	-0.015
.010	.008	.002	.022	- .012
.025	.021	.004	.038	- .012
.050	.032	.018	.055	- .005
.100	.065	.035	.100	.000
.900	.870	.030	.893	.007
.950	.927	.023	.943	.007
.975	.957	.018	.963	.012
.990	.973	.017	.978	.012
0.995	0.979	0.016	0.990	0.005

TABLE 15

COMPARISON OF ACTUAL ($\hat{p}$) AND THEORETICAL (p) CUMULATIVE PROPORTIONS AT 10 POINTS OF THE *t*-DISTRIBUTION
INPUT PARAMETERS: $\rho(T_X,T_Y) = .00$
$\rho(X,X') = .90$
$\rho(Y,Y') = .80$

p	$\hat{p}$ $r(T_X,T_Y)$	$p - \hat{p}$ $r(T_X,T_Y)$	$\hat{p}$ $r^{7/5}(T_X,T_Y)$	$p - \hat{p}$ $r^{7/5}(T_X,T_Y)$
N = 15, d.f. = 14				
0.005	0.020	-0.015	0.004	0.001
.010	.031	- .021	.009	.001
.025	.049	- .024	.016	.009
.050	.068	- .018	.032	.018
.100	.115	- .015	.058	.042
.900	.910	- .010	.960	- .060
.950	.946	.004	.973	- .023
.975	.968	.007	.982	- .007
.990	.975	.015	.992	- .002
0.995	0.978	0.017	0.997	-0.002
N = 30, d.f. = 29				
0.005	0.011	-0.006	0.002	0.003
.010	.015	- .005	.004	.006
.025	.038	- .013	.007	.018
.050	.060	- .010	.015	.035
.100	.103	- .003	.041	.059
.900	.887	.013	.955	- .055
.950	.931	.019	.979	- .029
.975	.959	.016	.992	- .017
.990	.978	.012	.995	- .005
0.995	0.989	0.006	0.997	-0.002
N = 60, d.f. = 59				
0.005	0.007	-0.002	0.000	0.005
.010	.018	- .008	.001	.009
.025	.033	- .008	.004	.021
.050	.051	- .001	.014	.036
.100	.101	- .001	.037	.067
.900	.904	- .004	.953	- .053
.950	.935	.015	.984	- .034
.975	.960	.015	.996	- .021
.990	.982	.008	.999	- .009
0.995	0.990	0.005	0.999	-0.004

TABLE 16

COMPARISON OF ACTUAL ($\hat{p}$) AND THEORETICAL (p) CUMULATIVE PROPORTIONS AT 10 POINTS OF THE *t*-DISTRIBUTION
INPUT PARAMETERS: $\rho(T_X,T_Y) = .00$
$\rho(X,X') = .80$
$\rho(Y,Y') = .80$

p	$\hat{p}$ $r(T_X,T_Y)$	$p-\hat{p}$ $r(T_X,T_Y)$	$\hat{p}$ $r^{7/5}(T_X,T_Y)$	$p-\hat{p}$ $r^{7/5}(T_X,T_Y)$
N = 15, d.f. = 14				
0.005	0.016	-0.011	0.003	0.002
.010	.025	- .015	.006	.004
.025	.035	- .010	.013	.012
.050	.058	- .008	.025	.025
.100	.106	- .006	.040	.060
.900	.875	.025	.955	- .055
.950	.932	.018	.982	- .032
.975	.961	.014	.988	- .013
.990	.982	.008	.993	- .003
0.995	0.988	0.007	0.996	-0.001
N = 30, d.f. = 29				
0.005	0.012	-0.007	0.003	0.002
.010	.018	- .008	.005	.005
.025	.033	- .008	.011	.014
.050	.062	- .012	.017	.037
.100	.096	.004	.034	.066
.900	.885	.015	.954	- .054
.950	.936	.014	.985	- .035
.975	.959	.016	.995	- .020
.990	.985	.005	.998	- .008
0.995	0.991	0.004	1.000	-0.005
N = 60, d.f. = 59				
0.005	0.011	-0.006	0.000	0.005
.010	.014	- .004	.000	.010
.025	.025	.000	.003	.022
.050	.055	- .005	.013	.037
.100	.113	- .013	.026	.074
.900	.892	.008	.963	- .063
.950	.940	.010	.991	- .041
.975	.965	.010	.995	- .020
.990	.990	.000	.998	- .008
0.995	0.992	0.003	0.999	-0.004

TABLE 17

COMPARISON OF ACTUAL ($\hat{p}$) AND THEORETICAL (p) CUMULATIVE PROPORTIONS AT 10 POINTS OF THE *t*-DISTRIBUTION

INPUT PARAMETER: $\rho(T_X,T_Y) = .00$

$\rho(X,X') = .90$

$\rho(Y,Y') = .50$

p	$\hat{p}$ $r(T_X,T_Y)$	$p - \hat{p}$ $r(T_X,T_Y)$	$\hat{p}$ $r^{7/5}(T_X,T_Y)$	$p - \hat{p}$ $r^{7/5}(T_X,T_Y)$
N = 15, d.f. = 14				
0.005	0.014	-0.009	0.000	0.005
.010	.021	- .011	.003	.007
.025	.035	- .010	.008	.017
.050	.057	- .007	.022	.028
.100	.095	.005	.040	.060
.900	.907	- .007	.946	- .046
.950	.933	.017	.973	- .023
.975	.957	.018	.986	- .011
.990	.975	.015	.990	.000
0.995	0.983	0.012	0.996	-0.001
N = 30, d.f. = 29				
0.005	0.011	-0.006	0.002	0.003
.010	.017	- .007	.004	.006
.025	.028	- .003	.008	.017
.050	.053	- .003	.015	.035
.100	.107	- .007	.031	.069
.900	.895	.005	.964	- .064
.950	.941	.009	.985	- .035
.975	.968	.007	.990	- .015
.990	.984	.006	.998	- .008
0.995	0.989	0.006	0.999	-0.004
N = 60, d.f. = 59				
0.005	0.010	-0.005	0.000	0.005
.010	.015	- .005	.000	.010
.025	.033	- .008	.007	.018
.050	.062	- .012	.011	.039
.100	.098	.002	.034	.066
.900	.901	- .001	.967	- .067
.950	.953	- .003	.991	- .041
.975	.971	.004	.997	- .022
.990	.989	.001	.999	- .009
0.995	0.997	-0.002	0.999	-0.004

.950, .975, .990, and .995 percentile points are summarized in Table 18 for $\underline{r}(\underline{T}_X,\underline{T}_Y)$ and in Table 19 for $\underline{r}^{7/5}(\underline{T}_X,\underline{T}_Y)$ for each combination of the input parameters.

The following results are summarized from the data presented in Tables 3-19.

The direction of the difference between the theoretical and actual cumulative proportions of jackknife statistics which were at or below the 10 percentile points of comparison varied across the different values of $\underline{\rho}(\underline{T}_X,\underline{T}_Y)$. For example, with $\underline{\rho}(\underline{T}_X,\underline{T}_Y)$ = 1.00 and for each combination of $\underline{N}$ and pair of reliabilities, the actual cumulative proportions exceeded the theoretical proportions at the upper five percentile points for $\underline{r}^{7/5}(\underline{T}_X,\underline{T}_Y)$. At the lower five percentile points, the reverse was true (see Tables 3-5). In contrast, with $\underline{\rho}(\underline{T}_X,\underline{T}_Y)$ = .50 and for each combination of $\underline{N}$ and pair of reliabilities, the actual cumulative proportion of jackknife statistics obtained for $\underline{r}^{7/5}(\underline{T}_X,\underline{T}_Y)$ exceeded the theoretical proportion at the five lower percentile points (see Tables 12-14).

The jackknife was sensitive to changes in the values of the reliabilities for each combination of $\underline{\rho}(\underline{T}_X,\underline{T}_Y)$ and $\underline{N}$. In general, the solutions obtained for $\underline{\rho}(\underline{X},\underline{X}')$ = .90, $\underline{\rho}(\underline{Y},\underline{Y}')$ = .50 were superior to the solutions obtained with the two higher pairs of reliabilities. Comparison of this result with the results describing the shape of the distribution of $\underline{r}(\underline{T}_X,\underline{T}_Y)$ (see Table 1) suggests that the best performance of the jackknife will be obtained in those situations in which the sampling distribution of the statistic to be jackknifed is approximately normally distributed with an appreciable variance.

TABLE 18

MEAN OF ABSOLUTE DIFFERENCES BETWEEN THEORETICAL AND ACTUAL PROPORTIONS OF JACKKNIFE STATISTICS AT 10 PERCENTILE POINTS OF THE *t*-DISTRIBUTION STATISTIC JACKKNIFED: $r(T_X, T_Y)$

$\rho(T_X,T_Y)$	$\rho(X,X')$	$\rho(Y,Y')$	Mean Absolute Difference N = 15	N = 30	N = 60
1.00	0.90	0.80	0.0235	0.0182	0.0114
1.00	.80	.80	.0257	.0206	.0125
1.00	.90	.50	.0244	.0244	.0207
0.90	.90	.80	.0449	.0440	.0227
.90	.80	.80	.0296	.0331	.0230
.90	.90	.50	.0221	.0205	.0135
.80	.90	.80	.0431	.0351	.0247
.80	.80	.80	.0413	.0380	.0264
.80	.90	.50	.0155	.0134	.0150
.50	.90	.80	.0403	.0332	.0171
.50	.80	.80	.0359	.0248	.0200
.50	.90	.50	.0161	.0171	.0165
.00	.90	.80	.0146	.0103	.0067
.00	.80	.80	.0122	.0093	.0059
0.00	0.90	0.50	0.0111	0.0059	0.0043
Overall Mean Absolute Difference			0.0269	0.0231	0.0160
Overall Mean Absolute Difference, $\rho(T_X,T_Y) > .00$			0.031	0.0267	0.0196

Note.—The mean absolute difference equals the mean of the absolute differences at the .005, .010, .025, .050, .100, .900, .950, .975, .990, and .995 percentile points of the *t*-distribution with the correct degrees of freedom, i.e.,

$$\text{Mean Absolute Difference} = \frac{\sum_{i=1}^{10} |p - \hat{p}|}{10} .$$

The overall mean absolute difference equals the mean of the mean absolute differences across each combination of $\rho(T_X,T_Y)$, $\rho(X,X')$, and $\rho(Y,Y')$ within each value of *N*.

TABLE 19

MEAN OF ABSOLUTE DIFFERENCES BETWEEN THEORETICAL AND ACTUAL PROPORTIONS OF JACKKNIFE STATISTICS AT 10 PERCENTILE POINTS OF THE *t*-DISTRIBUTION
STATISTIC JACKKNIFED: $r/s(T_X,T_Y)$

$\rho(T_X,T_Y)$	$\rho(X,X')$	$\rho(Y,Y')$	Mean Absolute Difference N = 15	N = 30	N = 60
1.00	0.90	0.80	0.0232	0.0170	0.0113
1.00	.80	.80	.0254	.0203	.0126
1.00	.90	.50	.0216	.0211	.0217
0.90	.90	.80	.0417	.0418	.0208
.90	.80	.80	.0253	.0306	.0218
.90	.90	.50	.0188	.0189	.0126
.80	.90	.80	.0356	.0306	.0216
.80	.80	.80	.0333	.0319	.0219
.80	.90	.50	.0085	.0099	.0094
.50	.90	.80	.0308	.0283	.0095
.50	.80	.80	.0290	.0209	.0120
.50	.90	.50	.0133	.0173	.0087
.00	.90	.80	.0165	.0229	.0259
.00	.80	.80	.0207	.0246	.0284
0.00	0.90	0.50	0.0198	0.0256	0.0281
Overall Mean Absolute Difference			0.0242	0.0241	0.0178
Overall Mean Absolute Difference, $\rho(T_X,T_Y)>.00$			0.0255	0.0241	0.0153

Note.—The mean absolute difference equals the mean of the absolute differences at the .005, .010, .025, .050, .100, .900, .950, .975, .990, and .995 percentile points of the *t*-distribution with the correct degrees of freedom, i.e.,

$$\text{Mean Absolute Difference} = \frac{\sum_{i=1}^{10} |p - \hat{p}|}{10}.$$

The overall mean absolute difference equals the mean of the mean absolute difference across each combination of $\rho(T_X,T_Y)$, $\rho(X,X')$, and $\rho(Y,Y')$ within each value of N.

There was a general tendency for the performance of the jackknife to improve for each combination of $\rho(T_X,T_Y)$ and pair of reliabilities as the value of N increased. The overall performance of the jackknife on $r^{7/5}(T_X,T_Y)$ was somewhat superior to the overall performance of the jackknife on $r(T_X,T_Y)$ for $\rho(T_X,T_Y)$ greater than zero. With $\rho(T_X,T_Y) > .00$, the overall mean of the means of the absolute differences between the theoretical and actual cumulative proportions at the 10 percentile points for $r^{7/5}(T_X,T_Y)$ was less than the corresponding overall mean for $r(T_X,T_Y)$. For $r^{7/5}(T_X,T_Y)$ the overall mean absolute differences equalled .0255 with $N = 15$, .0241 with $N = 30$, and .0153 with $N = 60$. The corresponding differences for $r(T_X,T_Y)$ equalled .0310, .0267, and .0196. For $\rho(T_X,T_Y) = .00$, the jackknife on $r(T_X,T_Y)$ yielded a considerably superior solution (see Tables 18 and 19).

For some of the combinations of the input parameters, the proportions of jackknife statistics falling in the critical regions corresponding to a two-tailed hypothesis text were not equal. For example, for $\rho(T_X,T_Y) = .90$, $\rho(X,X') = .90$, $\rho(Y,Y') = .80$, and $N = 60$, the probability of a Type I error equalled .068 (compared to the nominal value of .050). This indicates that 68 out of the 1000 statistics fell in the region of rejection. But in this case, 57 out of the 68 values fell in the region above $\rho(T_X,T_Y) = .90$ and only 11 in the region below $\rho(T_X,T_Y) = .90$. Hence, the Type I errors were not evenly distributed between the acceptance of $\rho(T_X,T_Y) < .90$ and the acceptance of

$\rho(T_X, T_Y) > .90$. For other combinations of the input parameters the proportions of jackknife statistics were more evenly divided, but the actual proportions were sometimes less, and other times greater than the nominal proportions.

To gain further information on the acceptability of the jackknife procedure, the results of the jackknife on $r(T_X, T_Y)$ and $r^{7/5}(T_X, T_Y)$ with N = 30 and N = 60 were compared to the results using normal curve theory.[3] Formula (3) was used to compute the standard error of $r(T_X, T_Y)$. The corresponding mean absolute differences for both the jackknife and the normal curve procedure are summarized in Tables 20 and 21. With N = 30, the solution obtained using normal curve theory was, in general, inferior to the jackknife solutions on $r(T_X, T_Y)$ and on $r^{7/5}(T_X, T_Y)$. The overall means of the absolute differences between the theoretical and actual cumulative proportions at the 10 points of comparison across the 15 combinations of $\rho(T_X, T_Y)$, $\rho(X, X')$, and $\rho(Y, Y')$ were .0231 for the jackknife on $r(T_X, T_Y)$, .0241 for the jackknife on $r^{7/5}(T_X, T_Y)$, and .0262 for the normal curve procedure. With N = 60, there was closer agreement between the three overall solutions. The overall means of the absolute difference were in this case .0160 for the jackknife on $r(T_X, T_Y)$, .0178 for the jackknife on $r^{7/5}(T_X, T_Y)$, and .0166 for the normal curve procedure.

3 Normal curve procedures were not employed for N = 15. Preliminary tests on the utility of equation (3) revealed that, with N = 15, the mean values for the standard error obtained from equation (3) were generally quite different from the observed standard deviation of the sampling distribution of $r(T_X, T_Y)$.

TABLE 20

MEAN OF ABSOLUTE DIFFERENCES BETWEEN THEORETICAL AND ACTUAL CUMULATIVE PROPORTIONS OF JACKKNIFE STATISTICS AT 10 PERCENTILE POINTS OF THE t-DISTRIBUTION AND MEAN OF ABSOLUTE DIFFERENCES BETWEEN THEORETICAL AND ACTUAL CUMULATIVE PROPORTIONS OF z-STATISTIC AT 10 PERCENTILE POINTS OF THE UNIT NORMAL N = 30

$\rho(T_X,T_Y)$	$\rho(X,X')$	$\rho(Y,Y')$	Mean Absolute Difference: Jackknife on $r^{7/5}(T_X,T_Y)$	Jackknife on $r(T_X,T_Y)$	Normal Curve
1.00	0.90	0.80	0.0170	0.0182	0.0279
1.00	.80	.80	.0203	.0206	.0180
1.00	.90	.50	.0211	.0224	.0256
0.90	.90	.80	.0418	.0440	.0476
.90	.80	.80	.0306	.0331	.0370
.90	.90	.50	.0189	.0205	.0165
.80	.90	.80	.0306	.0351	.0362
.80	.80	.80	.0319	.0380	.0408
.80	.90	.50	.0099	.0134	.0175
.50	.90	.80	.0283	.0332	.0351
.80	.80	.80	.0209	.0248	.0288
.50	.90	.50	.0173	.0171	.0217
.00	.90	.80	.0229	.0103	.0157
.00	.80	.80	.0246	.0093	.0129
0.00	0.90	0.50	0.0256	0.0059	0.0121
Overall Mean Absolute Difference			0.0241	0.0231	0.0262

Note.—The mean absolute difference equals the mean of the absolute differences at the .005, .010, .025, .050, .100, .900, .950, .975, .990, and .995 percentile points of the t-distribution with the correct degrees of freedom, i.e.,

$$\text{Mean Absolute Difference} = \frac{\sum_{i=1}^{10} |p - \hat{p}|}{10} .$$

The overall mean absolute difference equals the mean of the mean absolute differences across each combination of $\rho(T_X,T_Y)$, $\rho(X,X')$, and $\rho(Y,Y')$ within each value of N.

TABLE 21

MEAN OF ABSOLUTE DIFFERENCES BETWEEN THEORETICAL AND ACTUAL CUMULATIVE PROPORTIONS OF JACKKNIFE STATISTICS AT 10 PERCENTILE POINTS OF THE t-DISTRIBUTION AND MEAN OF ABSOLUTE DIFFERENCES BETWEEN THEORETICAL AND ACTUAL CUMULATIVE PROPORTIONS OF z-STATISTIC AT 10 PERCENTILE POINTS OF THE UNIT NORMAL N = 60

$\rho(T_X,T_Y)$	$\rho(X,X')$	$\rho(Y,Y')$	Mean Absolute Difference: Jackknife on $r^{7/5}(T_X,T_Y)$	Jackknife on $r(T_X,T_Y)$	Normal Curve
1.00	0.90	0.80	0.0113	0.0114	0.0074
1.00	.80	.80	.0126	.0125	.0110
1.00	.90	.50	.0217	.0207	.0172
0.90	.90	.80	.0208	.0227	.0256
.90	.80	.80	.0218	.0230	.0244
.90	.90	.50	.0126	.0135	.0108
.80	.90	.80	.0216	.0247	.0254
.80	.80	.80	.0219	.0264	.0255
.80	.90	.50	.0094	.0150	.0193
.50	.90	.80	.0095	.0171	.0182
.50	.80	.80	.0120	.0200	.0194
.50	.90	.50	.0087	.0165	.0209
.00	.90	.80	.0259	.0067	.0086
.00	.80	.80	.0284	.0059	.0079
0.00	0.90	0.50	0.0281	0.0043	0.0069
Overall Mean Absolute Difference			0.0178	0.0160	0.0166

Conclusions

1. The complexity of the sampling distribution of the estimate of the disattenuated correlation coefficient indicates that it is unlikely that a mathematical model can be formulated to describe the sampling distribution of $\underline{r}(\underline{T}_X, \underline{T}_Y)$.
2. The performance of the jackknife on the disattenuated correlation coefficient was sensitive to changes in the values of the input parameters. The performance was better for those combinations of the input parameters which yielded a sampling distribution of the statistic to be jackknifed that was more normally distributed and which had an appreciable variance.
3. The performance of the jackknife on $\underline{r}^{7/5}(\underline{T}_X, \underline{T}_Y)$ was slightly superior to the performance of the jackknife on $\underline{r}(\underline{T}_X, \underline{T}_Y)$ greater than zero. However, the discrepancies between the two performances became negligible as $\underline{N}$ increased.
4. The jackknife can be used to set approximate confidence intervals about $\underline{\rho}(\underline{T}_X, \underline{T}_Y)$, thereby communicating a general idea of the precision of the estimate obtained. However, the jackknife should not be used to perform directional hypotheses tests.
5. The jackknife fared well in comparison with the normal curve procedure. In the case of small $\underline{N}$ ($\underline{N}$ = 30), the jackknife was superior to the normal curve procedure.

The jackknife was originally proposed as a generally applicable procedure to construct approximate confidence intervals. The results obtained in this research support this claim for the disattenuated correlation coefficient, particularly for moderate size $\underline{N}$($\underline{N} \geq 30$). The jackknife is conceptually simple. Alternate procedures require the derivation of complex formulas for determining an adequate estimate of the standard error (vis., equation (3)). Current trends are for computer calculations to become less expensive and programming more expensive. If this trend continues, the advantages of first checking the jackknife to provide an estimate of the precision of a particular statistic, rather than struggling with the derivation of approximations of standard errors and hoping for normality, should become even more appealing as a method for communicating the precision of a given statistic. For situations in which the sample size is quite large, division of the sample into sub-samples larger than one, should reduce the computational time while at the same time retaining an acceptable degree of accuracy.

REFERENCES

Block, Jack. The equivalence of measures and the correction for attenuation. Psychological Bulletin 60:152-156; 1963.

Brillinger, D. R. The asympototic behavior of Tukey's general method of setting approximate confidence limits (the jackknife) when applied to maximum likelihood estimates. Review of the International Statistical Institute 32:202-206; 1964.

Brillinger, D. R. The application of the jackknife to the analysis of sample surveys. Commentary 8:74-80; 1966.

Cochran, W. G. Sampling Techniques. New York: John Wiley and Sons, 1963.

Cochran, W.G. Some effects of errors of measurement on multiple correlations. Journal of the American Statistical Association 65:22-35; 1970.

Collins, James R. Jackknifing generalizability. Unpublished Ph.D. Thesis, University of Colorado, 1970.

Cureton, Edward, E. Reliability and validity: basic assumptions and experimental designs. Educational and Psychological Measurement 25:327-346; 1965.

Cureton, Edward E. The stability coefficient. Educational and Psychological Measurement 31:45-55; 1971.

Cureton, Edward E. and Jack W. Dunlap. Spearman's correction for attenuation and its probable error. American Journal of Psychology 42:235-245; 1930.

Dempster, A. P. Estimation in multivariate analysis. In Paruchuri R. Kreshnaiah (ed.), Multivariate Analysis. New York, New York: Academic Press, Inc., 1966.

DuBois, Philip H. An Introduction to Psychological Statistics. New York, Harper & Row, 1965.

Forsyth, Robert A. and Leonard S. Feldt. An investigation of empirical sampling distributions of correlation coefficients corrected for attenuation. Educational and Psychological Measurement 29:61-71; 1969.

Forsyth, Robert A. An investigation of empirical sampling distributions of correlation coefficients corrected for attenuation. Unpublished Ph.D. Thesis, University of Iowa, 1967. (Available from University Microfilms, Inc., Ann Arbor; 67-9055).

Frankel, Martin R. Computing sampling errors for population estimates. Paper presented at the meeting of the International Association for Statistics in the Physical Sciences and the East-West Population Institute, Honolulu, August 1971.

Ghiselli, Edwin E. Theory of Psychological Measurement. New York: McGraw-Hill Book Company, 1964.

Gulliksen, Harold. Theory of Mental Tests. New York: John Wiley and Sons, 1950.

Hsu, Tse-Chi and Leonard S. Feldt. The effect of limitations on the number of criterion score values on the significance level of the F-test. American Educational Research Journal 6:515-527; 1969.

Johnson, Helmer G. Test reliability and correction for attenuation. Psychometrika 15:115-119; 1950.

Jones, Howard L. The jackknife method. In Proceedings of the I.B.M. Scientific Computing Symposium on Statistics, October 21-23, 1963. White Plains, New York: I.B.M. Data Processing Division, 1965.

Kelley, T. L. Statistical Method. New York: The Macmillan Company, 1923.

Kelley, T. L. Fundamentals of Statistics. Cambridge: Harvard University Press, 1947.

Lindquist, E. F. Design and Analysis of Experiments in Psychology and Education. Boston: Mifflin, 1953.

Lehman, R. S., and D. E. Bailey. Digital Computing: Fortran IV and Its Applications in Behavioral Science. New York: John Wiley and Sons, 1968.

Lord, Frederick M. A significance test for the hypothesis that two variables measure the same trait except for errors of measurement. Psychometrika 22:207-220; 1957.

Lord, Frederick M. Problems arising from the unreliability of the measuring instrument. In Philip H. DuBois and G. Douglas Mayo (eds.), Research Strategies for Evaluating Training. AERA Monograph Series on Curriculum Evaluation, No. 4. Chicago: Rand McNally & Co., 1970.

Lord, Frederick M. and Melvin R. Novick. Statistical Theories of Mental Test Scores. Reading, Massachusetts: Addison-Wesley, 1968.

Massey, Frank J., Jr. The Kolmogorov-Smirnov test for goodness of fit. Journal of American Statistical Association 46:68-78; 1951.

McNemar, Quinn. Attenuation and interaction. Psychometrika 23:259-265, 1958.

Miller, Rupert G. A trustworthy jackknife. Annals of Mathematical Statistics 35:1594-1605; 1964.

Miller, Rupert G. Jackknifing variances. Annals of Mathematical Statistics 39:567-582; 1968.

Mosteller, F., and J. W. Tukey. Data analysis, including statistics. In G. Lindsey and E. Aronson (eds.), Handbook of Social Psychology. (2nd ed.). Reading, Massachusetts: Addison-Wesley, 1968.

National Assessment of Educational Progress. Report 1: Science, National Results. National Assessment of Educational Progress, 1970.

Naylor, T. H., J. L. Balintfy, D. S. Burdick, and K. Chu. Computer Simulation Techniques. New York: Wiley, 1966.

Quenouille, M. H. Notes on bias in estimation. Biometrika 43:353-360, 1956.

Rogers, William Todd. Jackknifing disattenuated correlations. Unpublished Ph.D. Thesis, University of Colorado, 1971.

Spearman C. The proof and measurement of association between two things. American Journal of Psychology 15:72-101; 1904a.

Spearman C. General intelligence objectively determined and measured. American Journal of Psychology 15:201-293; 1904b.

Spearman C. Demonstration of formulae for true measurement of correlation. American Journal of Psychology 18:161-169; 1907.

Spearman C. Coefficient of correlation calculated from faulty data. British Journal of Psychology 3:271-295; 1910.

Shen, E. The standard error of certain estimated coefficients of correlation. Journal of Educational Psychology 15:462-465; 1924.

Thouless, Robert H. The effects of errors of measurement on correlation coefficients. British Journal of Psychology 29:383-403; 1939.

Tukey, J. W. Bias and confidence in not quite large samples. Annals of Mathematical Statistics 29:614; 1958.